Development and Characterization of EST-SSR Markers for Catalpa bungei (Bignoniaceae)

Authors: Wang, Peng, Ma, Yuzhu, Ma, Lingling, Li, Ya, Wang, Shu'an, et al.

Source: Applications in Plant Sciences, 4(4)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1500117

Applications in Plant Sciences 2016 4(4): 1500117

PRIMER NOTE

DEVELOPMENT AND CHARACTERIZATION OF EST-SSR MARKERS FOR *CATALPA BUNGEI* (BIGNONIACEAE)[1]

PENG WANG[2], YUZHU MA[2], LINGLING MA[2], YA LI[2,3], SHU'AN WANG[2], LINFANG LI[2], RUTONG YANG[2], AND QING WANG[2]

[2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing 210014, Jiangsu Province, People's Republic of China

- *Premise of the study: Catalpa bungei* (Bignoniaceae) is a deciduous tree native to China. We developed microsatellite markers for *C. bungei* to investigate its population genetics.
- *Methods and Results:* One hundred seventy-seven expressed sequence tag (EST)–simple sequence repeat (SSR) primer pairs were isolated and characterized using next-generation sequencing. Thirty of these primer pairs were polymorphic loci in 52 individuals of *C. bungei*. The number of alleles ranged from two to 18 with observed and expected heterozygosity values of 0.05–1.00 and 0.18–0.95, respectively. The fixation index ranged from –1.00 to 1.00 with an average of 0.32. No linkage disequilibrium was detected in any pair of loci. All markers showed good amplification results in four species (*C. bungei*, *C. fargesii*, *C. duclouxii*, and *C. ovata*) except three loci.
- *Conclusions:* These polymorphic markers are expected to be helpful in further studies on the systematics and phylogeography of *C. bungei* and related species.

Key words: Bignoniaceae; *Catalpa bungei*; expressed sequence tags; population genetics; RNA-seq; simple sequence repeats (SSRs).

Catalpa Scop. (Bignoniaceae) comprises 11 species of trees, and five of the 11 species in the genus originated in China. *Catalpa ovata* G. Don is distributed in central and northern China, whereas *C. bungei* C. A. Mey. and *C. fargesii* Bureau are distributed in central to southwestern China; *C. fargesii* has a glabrous form, namely, *C. duclouxii* Dode (Gilmour, 1936). *Catalpa tibetica* Forrest is endemic to southwestern China and, like *C. ovata*, has creamy yellow flowers. *Catalpa bungei* is characterized as fast growing, having excellent wood qualities, and being highly adaptable in China (Shi et al., 2011). Due to these economic and ecological benefits, it has been introduced and cultivated in Shandong, Jiangsu, Henan, and Anhui provinces (Shi et al., 2011). Molecular genetic studies have been few in number (Li, 2008), and no simple sequence repeats (SSRs) have been reported. To optimize the conservation and utilization of *C. bungei* and related species, the development of expressed sequence tag (EST)–SSR markers is very useful for germplasm identification and research into the genetic diversity of *C. bungei* and related species.

Next-generation sequencing (NGS) technologies have emerged as powerful tools for high-throughput EST sequence determination (Clark et al., 2013). EST-SSRs derived from EST sequences are more convenient and can be isolated with higher efficiency and at lower expense than genomic sequence SSRs (Wang et al., 2012). In this study, we identified 3999 SSR loci and characterized 30 polymorphic EST-SSR markers to facilitate our further investigations of systematics and population genetics in *C. bungei* and related species.

METHODS AND RESULTS

ESTs are an important source for the development of SSR markers. In this study, ESTs were isolated using a NGS approach. Total RNAs were extracted from the roots of one individual of *C. bungei* 'YU-1' using Trizol reagent according to the manufacturer's instructions (Invitrogen, Carlsbad, California, USA). Paired-end libraries with approximate average insert lengths of 200 bp were synthesized using a Genomic Sample Prep Kit (Illumina, San Diego, California, USA) according to the manufacturer's instructions. Libraries were sequenced (101-bp paired-end reads) on an Illumina HiSeq 2000 instrument by a customer sequencing service (Biomarker Technologies, Beijing, China). Raw reads were cleaned by removing adapter sequences, empty reads, and low-quality sequences. Clean reads were assembled into nonredundant transcripts using Trinity, which has been developed specifically for de novo assembly of transcriptomes using short reads (Grabherr et al., 2011). The clean sequence data has been deposited in the Short Read Archive database of the National Center for Biotechnology Information (NCBI; accession no. SRP059272). A total of 62,955 unigenes were obtained with an N50 length of 1417 bp. Potential SSR loci of these unigenes were detected using the MISA tool (Thiel et al., 2003; http://pgrc.ipk-gatersleben.de/misa). The parameters were as follows: minimum SSR motif length of 10 bp and repeat length of 10 for mononucleotides, six for dinucleotides, and five for tri-, tetra-, penta-, and hexanucleotides (Yang et al., 2014). A total of 3999 SSR loci were identified in 14,634 unigenes from the *C. bungei* transcriptome. Of these unigenes, 580 contained more than one SSR locus, and 484 SSR loci were present in compound formation. The combined set of all of the EST-SSR loci revealed that, on average, one EST-SSR was found for every 7.51 kb of sequence data. Within the identified EST-SSR loci, mono-, di-, tri-, tetra-, and pentanucleotide repeats had two, four, 10, 14, and two types, respectively. The most frequent repeat motifs were mononucleotide

[1]Manuscript received 19 October 2015; revision accepted 18 November 2015.

This work was supported by the Public Service Platform, Jiangsu Science and Technology Department (grant no. BM2012058), and the National Natural Science Foundation of China (grant no. 31200509).

[3]Author for correspondence: yalicnbg@icloud.com

doi:10.3732/apps.1500117

repeats (1957 [48.94%]), followed by dinucleotide (1164 [29.11%]), trinucleotide (834 [20.86%]), tetranucleotide (41 [1.02%]), and pentanucleotide repeats (3 [0.07%]) (Table 1). All of the dinucleotide and trinucleotide repeat motifs were further analyzed to determine their distribution. The most common dinucleotide motif was AG/CT (730 [62.71%]), and the rarest was CG/CG (5 [0.43%]) (Table 2). Among the trinucleotide repeats, AAG/CTT (243 [29.14%]) was the most common motif, followed by ATC/ATG (132 [15.83%]); ACT/AGT (9 [1.08%]) was the rarest motif (Table 2).

Subsequently, the mononucleotide repeats were discarded because it was difficult to distinguish genuine mononucleotide repeats from polyadenylation products and some were likely generated by base mismatching or sequencing errors. Primer pairs were designed using Primer3 (Rozen and Skaletsky, 1999). The major parameters for primer pair design were set as follows: primer length of 18–22 bases (optimal 20 bases), PCR product size of 100–500 bp (optimal 200 bp), GC content of 40–70% (optimal 50%), and annealing temperatures of 52–59°C (optimal 55°C). Based on these parameters, 177 primer pairs were designed and synthesized for polymorphism detection.

Genomic DNAs of all accessions were extracted from the leaves using a modified version of the cetyltrimethylammonium bromide (CTAB) method (Kabelka et al., 2002). Samples of *C. bungei* were collected from four populations: Luoning, Henan Province (population HN: 34°24′6″N, 111°42′42″E; n = 21); Chuxian, Anhui Province (population AH: 32°50′54″N, 117°47′49″E; n = 11); Lianyungang, Jiangsu Province (population JS: 34°40′3″N, 119°19′60″E; n = 6); Qingzhou, Shandong Province (population SD: 36°46′15″N, 118°25′56″E; n = 14). Samples of three related species were collected from three populations: *C. duclouxii* in Kunming, Yunnan Province (25°02′32″N, 102°38′46″E; n = 13); *C. fargesii* in Yishui, Shandong Province (35°48′38″N, 118°38′5″E; n = 15); and *C. ovata* in Yunxian, Hubei Province (32°51′33″N, 110°44′10″E; n = 12). Plants for all accessions were grown in the *Catalpa* germplasm repository at the Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, and vouchers are deposited at the Herbarium of the Institute of Botany, Jiangsu Province and Chinese Academy of Sciences (NAS), Nanjing, China (Appendix 1). Approximately 10 g of young leaves were collected in the spring season. PCR amplification was carried out in 10-μL reaction mixtures containing 30 ng of template DNA, 1× PCR buffer (Mg^{2+} free), 2.0 mM $MgCl_2$, 0.2 mM dNTPs, 0.25 μM of each primer, and 1 unit *Taq* polymerase (TaKaRa Biotechnology Co., Dalian, China). Cycling was performed on a T100 Thermal Cycler (Bio-Rad, Marnes-la-Coquette, France). Amplification reactions were initiated with a pre-denaturing step (95°C for 5 min), followed by denaturing (95°C for 30 s), annealing (55°C for 45 s), extension (72°C for 60 s) for 32 cycles, and a final extension at 72°C for 8 min. Amplified PCR products were separated on 8% denaturing polyacrylamide gels using a vertical electrophoresis device. Detection of EST-SSR bands was performed using the silver staining method.

One hundred seventy-seven EST-SSR primer pairs were synthesized in this study. Fifty-five primer pairs were identified that yielded stable, clear, and repeatable amplicons in all accessions. The other primer pairs were unstable or gave no product. The 55 primers corresponded to 25 monomorphic loci (Appendix S1) and 30 polymorphic loci (Table 3). The polymorphic SSR loci were analyzed with POPGENE version 1.32 software (Yeh et al., 1999) for the number of alleles per locus (A), observed heterozygosity (H_o), expected heterozygosity (H_e), and fixation index (F_{IS}). The A values ranged from two to 18 with a mean of 6.78 (Table 4). The H_o and H_e values were 0.05–1.00 and 0.18–0.95 with averages of 0.53 and 0.75, respectively. The F_{IS} values ranged from –1.00 to 1.00 with an average of 0.32. Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium using Bonferroni correction were tested for every locus. Less than half of the loci (12, six, one, and seven loci in populations HN, AH, JS, and SD, respectively) showed significant departure from HWE ($P < 0.001$). Significant linkage disequilibrium was not detected between any pair of loci ($P < 0.001$).

Table 1. EST-SSRs present in the *Catalpa bungei* transcriptome.

Repeat type	No. of motif types	No. of EST-SSRs	Proportion in all SSRs (%)
Mononucleotide	2	1957	48.937
Dinucleotide	4	1164	29.107
Trinucleotide	10	834	20.855
Tetranucleotide	14	41	1.025
Pentanucleotide	2	3	0.075
Total	32	3999	100

Table 2. Characteristics of the di- and trinucleotide repeat motifs in the *Catalpa bungei* transcriptome.

Repeat type	Repeat motif	Number	Proportion (%)
Dinucleotide	AC/GT	217	18.64
	AG/CT	730	62.71
	AT/AT	212	18.21
	CG/CG	5	0.43
Trinucleotide	AAC/GTT	33	3.96
	AAG/CTT	243	29.14
	AAT/ATT	63	7.55
	ACC/GGT	91	10.91
	ACG/CGT	11	1.32
	ACT/AGT	9	1.08
	AGC/CTG	91	10.91
	AGG/CCT	70	8.39
	ATC/ATG	132	15.83
	CCG/CGG	91	10.91

Cross-amplification of 30 polymorphic loci was tested in 61 individuals of four *Catalpa* species under the same PCR conditions used for *C. bungei*. All markers showed successful amplification results in more than half of the 61 individuals tested, with the exception of three loci (comp100847, comp111793, and comp114074) (Table 5).

To identify potential functions of the 30 polymorphic SSR-associated unigenes, the sequences were used for BLAST searches and annotation against the NCBI nonredundant protein (NR) database (http://www.ncbi.nlm.nih.gov/) using an E-value cut-off of 10^{-5}. All sequences were found to have potential functions by BLASTX. These sequences showed significant homology to protein sequences from *Sesamum indicum* L., *Rehmannia glutinosa* (Gaertn.) Libosch. ex Fisch. & C. A. Mey., *Genlisea aurea* A. St.-Hil., and *Erythranthe guttata* (DC.) G. L. Nesom. The potential functions were mainly related to transcription factor, hormone metabolism, and carbon metabolism (Appendix S2).

CONCLUSIONS

In the present study, we have developed 30 novel EST-SSR polymorphic markers for *C. bungei*. These markers provide an efficient tool for investigating population genetic diversity in different environments and will facilitate studies on molecular breeding, genetic improvement, and conservation of *C. bungei* and related species.

LITERATURE CITED

Clark, S. M., V. Vaitheeswaran, S. J. Ambrose, R. W. Purves, and J. E. Page. 2013. Transcriptome analysis of bitter acid biosynthesis and precursor pathways in hop (*Humulus lupulus*). *BMC Plant Biology* 13: 12.

Gilmour, J. S. L. 1936. *Catalpa fargesii* forma *duclouxii*. *Curtis's Botanical Magazine* 159: tab. 9458.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644–652.

Kabelka, E., B. Franchino, and D. M. Francis. 2002. Two loci from *Lycopersicon hirsutum* LA407 confer resistance to strains of *Clavibacter michiganensis* subsp. *michiganensis*. *Phytopathology* 92: 504–510.

Li, J. H. 2008. Phylogeny of *Catalpa* (Bignoniaceae) inferred from sequences of chloroplast *ndhF* and nuclear ribosomal DNA. *Journal of Systematics and Evolution* 46: 341–348.

Rozen, S., and H. Skaletsky. 1999. Primer3 on the WWW for general users and for biologist programmers. *In* S. Misener and S. A. Krawetz [eds.], Methods in molecular biology, vol. 132: Bioinformatics methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

TABLE 3. Characteristics of 30 polymorphic EST-SSR markers in *Catalpa bungei*.[a]

Locus	Primer sequences (5′–3′)	Repeat motif	Product size (bp)	GenBank accession no.
comp100219	F: CAGGGAGTTTTCCGATTCAA	$(CAG)_5$	258	KT893751
	R: TTTGCCCGTATTTTCTCCAG			
comp100274	F: CGCTTTACCATTTGAAGGGA	$(GA)_6$	256	KT893752
	R: AACATCAGGCTTAACGGTGG			
comp100480	F: CGTGGATGAACACGAACAAC	$(TCC)_5$	210	KT893753
	R: CCTTTCCTTCTTCCACCTCC			
comp100607	F: ACCGGAGCAAAACAAAAATG	$(TCC)_5$	189	KT893754
	R: AAAGCCCGTACTCTTGCTGA			
comp100745	F: GCAAAATCCGTTGTTTCGAT	$(TG)_8$	278	KT893755
	R: GCAAGAGGCAAGACATAGCC			
comp100817	F: AAAGCAGCACGAGACGAGAT	$(TTG)_6$	183	KT893756
	R: ATTACAACATTTGCAGGGGC			
comp100847	F: CTGCCTCCTGCATTCTCTTC	$(GGC)_5$	186	KT893757
	R: GTTCGGGATCGTCGTCTTTA			
comp102534	F: AGAAAGCCCAATTGCAACAC	$(ATC)_5$	215	KT893758
	R: ACAACCACTCAACTCTGGGG			
comp103435	F: CTTTCCCCCATTGTTTGTTG	$(TAT)_5$	149	KT893759
	R: AAATGGTGGACAGGATTGGA			
comp104537	F: ACCGCTTGCATCTCTGATTT	$(GAG)_5$	263	KT893760
	R: CACTCATCTCTCACCACCGA			
comp107379	F: GTACTGCACCCACCCTCCTA	$(CAC)_5$	191	KT893761
	R: GGGGAAGAGGGTTGAAAGAG			
comp108461	F: CAAACGAATGTACGGAGGCT	$(GCG)_6$	168	KT893762
	R: GAGATTGAGGCAGAGGATGC			
comp108487	F: CAACAGCAATGAATGTTGGG	$(ATG)_5$	141	KT893763
	R: TCGTAGGCGGTCCATAACTC			
comp109601	F: TCCTGCTAATTTACCCAGAACAA	$(AT)_9$	155	KT893764
	R: GAACGTAGCGAGAAGAACCG			
comp109989	F: GCTTATGGGGGTCAAATCAA	$(ATT)_5$	127	KT893765
	R: GCATGAGAGGGAGACCAGAG			
comp110079	F: TGCAGAGTGGATCAAGCAAG	$(AT)_9$	215	KT893766
	R: CTACAAAACCCTGCGCGTAT			
comp110536	F: GCTCGCCTGTCAAGAAAATC	$(TGG)_5$	253	KT893767
	R: GCTCTTGTTCCAAAGGCTTG			
comp110884	F: ATGACACCCATCTTCTTCGC	$(GTC)_8$	258	KT893768
	R: GGAAGATGAGAGCAAGCCAG			
comp111793	F: CGATTTTCCAGAGGGAATGA	$(CGG)_6$	273	KT893769
	R: CTATGCTACTTCGCCGAACC			
comp112144	F: CCCTCTGTTCACTCCCGATA	$(AC)_9$	203	KT893770
	R: GGAGAAGTGCAAAGGAGACG			
comp112643	F: ACCAGGTGGCTGAATGAATC	$(CA)_7$	278	KT893771
	R: GCAGGCACAAAATCATGAGA			
comp112777	F: CTTCTGGAATCCTCCCTTCC	$(TC)_7$	227	KT893772
	R: GAATCGAAGGAGACTGCGAC			
comp112944	F: AATGCTTATAATGCCAGCGG	$(CCT)_5$	115	KT893773
	R: GCCTCACAACAGCAAGTTCA			
comp112997	F: AAAACGACAGCAACCAAACC	$(AGA)_5$	263	KT893774
	R: AAAGAGTTGGACAAGCCGAA			
comp113774	F: CCTGAAGCCTAATCTGCCTG	$(CCG)_5$	251	KT893775
	R: TGAAGTTGAAAATGGAGCCC			
comp113869	F: TGGCGGCTCTCCATATTAAC	$(TTC)_5$	113	KT893776
	R: ATGCAGGGCAGAGACAGAGT			
comp113985	F: CTTGGAGCGACGTTTCTTTC	$(CAT)_5$	176	KT893777
	R: CCAAACATCATCGACAAACG			
comp114074	F: GTGGTGGGCACTCCTTAGAA	$(AG)_7$	269	KT893778
	R: GCAACAGGCCAATACATCCT			
comp114135	F: CTGGCTTCCGAAATTGTGTT	$(TGA)_6$	213	KT893779
	R: TCATCATCCTGTGATGCGAT			
comp114163	F: CGCTCTCTTCAAGCTGCTCT	$(TC)_6$	103	KT893780
	R: GATGATGAATCCGAGGAGGA			

[a]Annealing temperature for all loci was 55°C.

SHI, X., Y. LI, R. T. YANG, Q. WANG, AND G. YAO. 2011. ISSR analysis of genetic diversity of *Catalpa bungei* C.A.Mey. germplasm resources in China. *Jiangsu Journal of Agricultural Sciences* 3: 32.

THIEL, T., W. MICHALEK, R. K. VARSHNEY, AND A. GRANER. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 10: 411–422.

TABLE 4. Genetic properties of 30 polymorphic EST-SSR loci in *Catalpa bungei*.[a]

	Population HN ($N = 21$)				Population AH ($N = 11$)				Population JS ($N = 6$)				Population SD ($N = 14$)			
Locus	A	H_o	H_e	F_{IS}	A	H_o	H_e	F_{IS}	A	H_o	H_e	F_{IS}	A	H_o	H_e	F_{IS}
comp100219	7	0.38*	0.82	0.52	6	0.45	0.79	0.40	5	0.17*	0.80	0.77	8	0.21*	0.87	0.74
comp100274	8	0.57	0.80	0.27	9	0.73	0.89	0.15	5	0.67	0.85	0.14	7	0.36	0.85	0.57
comp100480	3	0.33*	0.62	0.45	3	0.18*	0.59	0.68	4	0.33	0.79	0.54	5	0.36	0.60	0.39
comp100607	9	0.43	0.79	0.44	9	0.45	0.87	0.45	6	0.17	0.86	0.79	13	0.64	0.90	0.26
comp100745	6	0.14	0.55	0.73	6	0.18	0.71	0.73	4	0.17	0.65	0.72	7	0.21*	0.82	0.73
comp100817	3	0.67*	0.61	−0.12	3	0.82	0.60	−0.42	2	1.00	0.55	−1.00	4	0.57*	0.73	0.19
comp100847	6	0.14*	0.78	0.81	4	0.27*	0.58	0.50	4	0.67	0.77	0.06	6	0.21	0.62	0.64
comp102534	4	0.00*	0.47	1.00	5	0.27*	0.79	0.64	7	0.33	0.88	0.59	7	0.50	0.83	0.37
comp103435	7	0.05*	0.65	0.92	5	0.18	0.58	0.67	5	0.17	0.80	0.77	9	0.50	0.79	0.34
comp104537	5	0.29*	0.72	0.59	7	0.45	0.82	0.42	6	0.67	0.85	0.14	8	0.93	0.84	−0.14
comp107379	3	0.48*	0.65	0.25	4	0.73*	0.74	−0.02	5	0.83	0.83	−0.09	4	0.57*	0.71	0.16
comp108461	9	0.48	0.77	0.36	8	0.64	0.87	0.23	4	0.17	0.56	0.68	8	0.79	0.86	0.05
comp108487	18	0.76	0.94	0.17	10	0.36	0.91	0.58	6	0.33	0.82	0.56	14	0.64	0.92	0.28
comp109601	13	0.67	0.87	0.22	14	0.82	0.95	0.10	9	0.67	0.95	0.24	9	0.79	0.84	0.03
comp109989	9	0.90	0.82	−0.14	8	0.82	0.88	0.03	5	0.83	0.80	−0.13	13	0.93	0.93	−0.03
comp110079	7	0.29	0.70	0.58	8	0.45	0.84	0.43	4	0.17	0.56	0.68	9	0.29	0.74	0.60
comp110536	10	0.62*	0.58	−0.09	8	0.55	0.75	0.23	7	0.74	0.78	0.46	10	0.79	0.84	0.23
comp110884	11	0.43	0.70	0.38	10	0.27*	0.85	0.33	6	0.78	0.85	0.13	13	0.64	0.82	0.46
comp111793	8	0.67	0.78	0.95	11	0.82	0.86	0.39	5	0.84	0.72	−0.36	14	0.64	0.58	−0.11
comp112144	9	0.57*	0.61	0.21	7	0.55	0.83	0.52	5	0.83	0.79	−0.28	8	0.36	0.86	0.79
comp112643	4	0.86	0.45	−0.32	6	0.88	0.82	−0.12	3	0.68	0.60	−0.08	6	0.79	0.79	0.03
comp112777	6	0.33*	0.69	0.45	5	0.57	0.18	−0.58	3	0.50	0.57	0.15	7	0.43	0.70	0.65
comp112944	5	0.86	0.62	0.33	6	0.73	0.75	0.34	4	0.71	0.75	0.14	3	0.79	0.60	−0.32
comp112997	3	0.63	0.53	0.29	4	0.73	0.76	0.21	2	0.81	0.86	0.21	4	0.50	0.80	0.63
comp113774	6	0.35*	0.95	0.85	7	0.82	0.85	0.13	4	0.74	0.75	0.16	5	0.50	0.76	0.52
comp113869	9	0.68	0.71	0.49	6	0.55	0.78	0.37	7	0.78	0.88	0.23	8	0.07*	0.81	0.76
comp113985	5	0.35	0.52	0.62	3	0.27*	0.71	0.79	4	0.76	0.81	0.27	6	0.43	0.68	0.38
comp114074	8	0.52	0.23	−0.25	5	0.36	0.83	0.64	5	0.51	0.83	0.43	7	0.21*	0.83	0.89
comp114135	12	0.24	0.76	0.39	8	0.64	0.87	0.42	9	0.75	0.87	0.13	10	0.36	0.89	0.78
comp114163	7	0.80	0.71	−0.26	9	0.80	0.55	−0.55	6	0.88	0.91	0.24	6	0.21*	0.81	0.69
Average	7.33	0.48	0.68	0.37	6.80	0.54	0.76	0.29	5.03	0.59	0.78	0.24	7.93	0.51	0.79	0.39

Note: A = number of alleles; F_{IS} = fixation index; H_e = expected heterozygosity; H_o = observed heterozygosity.

[a]Locality information for populations: HN = Luoning, Henan Province (34°24′6″N, 111°42′42″E); AH = Chuxian, Anhui Province (32°50′54″N, 117°47′49″E); JS = Lianyungang, Jiangsu Province (34°40′3″N, 119°19′60″E); SD = Qingzhou, Shandong Province (36°46′15″N, 118°25′56″E).

*Designates significant deviation from Hardy–Weinberg equilibrium genotypic proportions after sequential Bonferroni correction for multiple tests ($P < 0.01$).

TABLE 5. Cross-amplification results for the 30 polymorphic EST-SSR loci in 61 individuals of four *Catalpa* species.

ID[a]	M01[b]	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30
CB001	1	1	1	0	0	1	0	0	0	1	0	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	0	1	1
CB002	1	1	1	0	1	1	0	0	0	1	0	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1
CB003	1	1	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1
CB012	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	0	1	1
CB014	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	0	1	1
CB018	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
CB021	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1
CB031	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1
CB032	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	0	1	1	1	1	1	0	1	1
CB033	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	0	1	1
CB038	0	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
CB039	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1
CB040	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1
CB041	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
CB042	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	0	1	1	0	1	1	1	0	1	0	1	1
CB043	0	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1
CB044	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
CB048	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	0	1	1	0	1	1	1	0	1	0	1	1
CB060	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1
CB067	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0	1	1
CB069	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1
CD009	0	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1
CD013	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1
CD034	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1
CD050	0	1	1	1	0	1	0	0	0	1	0	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	0	1	1
CD053	1	1	1	1	1	1	0	0	0	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1
CD057	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	0	1	1
CD070	1	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1
CD073	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1
CD078	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1
CD097	0	0	0	0	0	1	0	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CD098	0	0	0	0	0	1	0	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CD099	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CD101	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF024	1	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF027	1	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF051	0	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF059	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF062	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF064	1	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF066	1	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF100	1	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF102	1	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF103	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF104	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF105	1	1	1	1	0	1	1	1	0	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF106	1	1	0	1	0	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF107	1	1	1	1	0	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CF108	1	1	1	1	0	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CO015	1	1	1	1	0	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CO019	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CO083	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CO086	0	1	1	1	1	1	0	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	1	0	1	1
CO092	0	0	1	0	1	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1	0	0	0	1	0	1	1

TABLE 5. Continued.

ID[a]	M01[b]	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30
CO093	0	1	1	0	1	1	0	0	0	1	1	1	1	1	1	0	1	0	0	0	1	0	1	1	0	1	1	0	1	1
CO094	0	1	0	0	1	0	0	1	0	1	0	1	1	1	0	1	1	0	0	1	0	0	1	0	0	0	1	0	1	1
CO095	0	1	0	0	1	0	0	0	1	1	0	1	1	1	0	0	0	0	1	0	1	0	1	0	0	0	1	0	1	1
CO096	0	1	0	0	1	0	0	1	0	1	0	1	1	1	0	0	0	0	0	0	1	0	1	0	0	0	1	0	1	1
CO063	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	0
CO004	0	1	0	1	0	1	0	1	0	1	1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	0	1	0	1	0
CO037	0	1	0	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	0

Note: 1 = successful amplification; 0 = failed amplification.

[a]Accession numbers/sample identification codes at the *Catalpa* germplasm repository, Institute of Botany, Jiangsu Province and Chinese Academy of Sciences. CB = *Catalpa bungei*; CD = *Catalpa duclouxii*; CF = *Catalpa fargesii*; CO = *Catalpa ovata*; M01 = locus comp100219; M02 = locus comp100274; M03 = locus comp100480; M04 = locus comp100607; M05 = locus comp100745; M06 = locus comp100817; M07 = locus comp100847; M08 = locus comp102534; M09 = locus comp103435; M10 = locus comp104537; M11 = locus comp107379; M12 = locus comp108461; M13 = locus comp108487; M14 = locus comp109601; M15 = locus comp109989; M16 = locus comp110079; M17 = locus comp110536; M18 = locus comp110884; M19 = locus comp111793; M20 = locus comp112144; M21 = locus comp112643; M22 = locus comp112777; M23 = locus comp112944; M24 = locus comp112997; M25 = locus comp113774; M26 = locus comp113869; M27 = locus comp113985; M28 = locus comp114074; M29 = locus comp114135; M30 = locus comp114163.

WANG, L. H., Y. X. ZHANG, X. Q. QI, Y. GAO, AND X. R. ZHANG. 2012. Development and characterization of 59 polymorphic cDNA-SSR markers for the edible oil crop *Sesamum indicum* (Pedaliaceae). *American Journal of Botany* 99: e394–e398.

YANG, T., J. JIANG, M. BURLYAEVA, J. HU, C. J. COYNE, S. KUMAR, R. REDDEN, ET AL. 2014. Large-scale microsatellite development in grasspea (*Lathyrus sativus* L.), an orphan legume of the arid areas. *BMC Plant Biology* 14: 65.

YEH, F. C., R. C. YANG, AND T. BOYLE. 1999. POPGENE version 1.32: Microsoft Windows-based freeware for population genetic analysis, quick user guide. Center for International Forestry Research, University of Alberta, Edmonton, Alberta, Canada.

Appendix 1. Voucher information for *Catalpa* species used in this study.

Species	Voucher specimen accession no.[a]	Collection locality[b]	Geographic coordinates	No. of individuals
C. bungei	CB-HN-2010-SX	Luoning, Henan Province	34°24′6″N, 111°42′42″E	21
C. bungei	CB-AH-2010-SX	Chuxian, Anhui Province	32°50′54″N, 117°47′49″E	11
C. bungei	CB-JS-2010-SX	Lianyungang, Jiangsu Province	34°40′3″N, 119°19′60″E	6
C. bungei	CB-SD-2010-SX	Qingzhou, Shandong Province	36°46′15″N, 118°25′56″E	14
C. duclouxii	CD-KM-2010-YG	Kunming, Yunnan Province	25°02′32″N, 102°38′46″E	13
C. fargesii	CF-YS-2010-YG	Yishui, Shandong Province	35°48′38″N, 118°38′5″E	15
C. ovata	CF-YX-2010-YG	Yunxian, Hubei Province	32°51′33″N, 110°44′10″E	12

Note: SX = Xin Shi, collector; YG = Gan Yao, collector.

[a] Vouchers deposited at the Herbarium of the Institute of Botany, Jiangsu Province and Chinese Academy of Sciences (NAS), Nanjing, China.

[b] Locality and Chinese province.